

A new method to break the mass sheet degeneracy using aperture moments

Markus Rexroth¹, Priyamvada Natarajan², & Jean-Paul Kneib^{1,3}

¹ *Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland*

² *Department of Astronomy, Yale University, 260 Whitney Avenue, New Haven, CT 06511, USA*

³ *Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Mass determinations from gravitational lensing shear and the higher order estimator flexion are both subject to the mass sheet degeneracy. Mass sheet degeneracy refers to a transformation that leaves the reduced shear and flexion invariant. In general, this transformation can be approximated by the addition of a constant surface mass density sheet. We propose a new technique to break the mass sheet degeneracy. The method uses mass moments of the shear or flexion fields in combination with convergence information derived from number counts which exploit the magnification bias. The difference between the measured mass moments provides an estimator for the magnitude of the additive constant that is the mass-sheet. For demonstrating this, we derive relations that hold true in general for n -th order moments and show how they can be employed effectively to break the degeneracy. We investigate the detectability of this degeneracy parameter from our method and find that the degeneracy parameter can be feasibly determined from stacked galaxy-galaxy lensing data and cluster lensing data. Furthermore, we compare the signal-to-noise ratios of convergence information from number counts with shear and flexion for SIS and NFW models. We find that the combination of shear and flexion performs best on galaxy and cluster scales and the convergence information can therefore be used to break the mass sheet degeneracy without quality loss in the mass reconstruction. In summary, there is power in the combination of shear, flexion, convergence and their higher order moments. With the anticipated wealth of lensing data from upcoming and future satellite missions - EUCLID and WFIRST - this technique will be feasible.

Key words: Gravitational lensing; weak – galaxies: haloes – galaxies: clusters: general – dark matter

1 INTRODUCTION

Many mass reconstructions were successfully obtained utilizing gravitational lensing shear (e.g., [Hoekstra et al. 2004](#); [Natarajan et al. 1998](#)) and the feasibility of lens models from flexion has been demonstrated ([Leonard et al. 2007](#); [Okura et al. 2008](#)). Several mass reconstruction methods have been developed (e.g., [Schneider et al. 2000](#); [Bartelmann et al. 1996](#); [Kaiser 1995](#); [Kaiser & Squires 1993](#)), see [Kneib & Natarajan \(2011\)](#) for a review of galaxy cluster lens reconstructions. However, the masses derived from the shear and flexion fields are only determined up to a constant due to the mass sheet degeneracy (e.g., [Schneider & Seitz 1995](#); [Schneider & Er 2008](#)). Several techniques have been proposed to lift the degeneracy. [Broadhurst et al. \(1995\)](#) suggested using magnification information by comparing lensed and unlensed background

source counts to reconstruct the non-degenerate mass sheet. This effect was subsequently observed and applied (e.g. [Fort et al. 1997](#); [Taylor et al. 1998](#); [Dye et al. 2002](#)). [Bartelmann & Narayan \(1995\)](#) proposed the lens parallax method. This method compares the mean sizes of lensed faint blue galaxies with those of unlensed sources in an empty control field. While gravitational lensing magnifies the area, it preserves the surface brightness, and thus the magnification can be inferred by comparing the mean sizes of lensed and unlensed galaxies with the same surface brightness. The magnification information can then be combined with the shear measurement to break the degeneracy. However, it is hard to estimate the surface brightness from seeing-convolved images and thus the application of this technique is challenging ([Meylan et al. 2006](#)). In addition, the degeneracy can also be lifted by fitting the

gravitational lens potential to shear and magnification data simultaneously. Several fitting and Bayesian methods have been proposed for this purpose (see e.g. Umetsu et al. 2011; Schneider et al. 2000; Bartelmann et al. 1996). Bradač et al. (2004) showed that it is possible to reconstruct non-degenerate mass distributions if the individual redshifts of sources at different distances are combined with the shear field. However, this technique can only be used for critical lenses. A different method is needed for sub-critical lenses and lenses where a sufficient number of redshifts is not available. In this paper, we present, develop, and explicate the technical framework that can be used to derive the mass sheet degeneracy using combinations of shape and convergence estimator moments.

We propose a new method that can break the mass-sheet degeneracy by a simple comparison of aperture masses. We show that taking the difference in mass estimates from convergence information, derived for example from source counts, and shear or flexion data allows the determination of the value of the mass sheet, thus enabling the derivation of calibrated masses. We demonstrate its feasibility for both stacked galaxy-galaxy lensing and cluster lensing. Furthermore, we show that for two widely used mass models, the signal-to-noise of shear and flexion in the weak lensing regime is always superior to the number counts, both on galaxy and cluster scales. Therefore we can use the convergence information to break the mass sheet without sacrificing the quality of the lens reconstruction.

We develop the theoretical framework of this method for both shear and flexion, the higher order lensing effect which describes the arciness of the lensed sources. Flexion was introduced into the weak lensing framework, because it provides valuable additional information and is highly sensitive to substructure (Bacon et al. 2006; Irwin & Shmakova 2006; Goldberg & Bacon 2005; Goldberg & Natarajan 2002). Studies have showed that the addition of flexion can improve mass reconstructions significantly (e.g., Leonard et al. 2007; Okura et al. 2007). The high quality data from the *Hubble Frontier Fields* initiative permit flexion measurements with increased accuracy (Rexroth 2015).

The paper is organized as follows. In section 2, we give a brief description of the weak lensing formalism and the mass-sheet degeneracy. Section 3 presents the equivalent moments and how their transformations can break the degeneracy. We demonstrate this method for a singular isothermal sphere (SIS) halo model in section 4. In section 5 we investigate the detectability of the mass sheet parameter without making assumptions on the halo model. Section 6 shows that the combination of shear and flexion has a higher signal-to-noise ratio than number counts for SIS and Navarro-Frenk-White (NFW) halos. We conclude in section 7.

2 WEAK LENSING FORMALISM AND THE MASS-SHEET DEGENERACY

Weak lensing of background sources by massive foreground objects results in the deformation of their shapes. The

strength of these distortions is directly related to the surface mass density of the lens. Let us consider a lens with projected surface mass density $\Sigma(\boldsymbol{\theta})$, where the angular coordinate $\boldsymbol{\theta}$ denotes the position in the lens plane. The convergence of the lens κ is defined as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{crit}}}, \quad (1)$$

where the critical surface mass density, defined as

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \left(\frac{D_{\text{os}}}{D_{\text{ls}} D_{\text{ol}}} \right), \quad (2)$$

depends on the angular diameter distances from the lens to the source D_{ls} , observer to source D_{os} , and observer to lens D_{ol} . The lensed and unlensed coordinates for the distorted background sources are related by the mapping

$$A_{ij}(\boldsymbol{\theta}) = \frac{\partial \theta'_i}{\partial \theta_j}, \quad (3)$$

$$A(\boldsymbol{\theta}) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (4)$$

where $\boldsymbol{\theta}'$ are the unlensed coordinates and γ is the shear. The magnification of the sources is

$$\mu = \frac{1}{(1 - \kappa)^2 - \gamma^2}. \quad (5)$$

Furthermore, we can express the convergence and shear using the deflection potential

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\beta \kappa(\boldsymbol{\beta}) \ln |\boldsymbol{\theta} - \boldsymbol{\beta}| \quad (6)$$

as

$$\kappa = \frac{1}{2}(\psi_{,11} + \psi_{,22}), \quad \gamma = \begin{pmatrix} \frac{1}{2}(\psi_{,11} - \psi_{,22}) \\ \psi_{,12} \end{pmatrix}. \quad (7)$$

The commas denote partial θ derivatives. The reduced shear is defined as

$$g = \frac{\gamma}{1 - \kappa}. \quad (8)$$

If shear and convergence are constant over a source image, the transformation is given by a simple matrix multiplication:

$$\theta'_i = A_{ij} \theta_j. \quad (9)$$

However, if the shear varies over the image, we have to expand to include flexion terms:

$$\theta'_i \simeq A_{ij} \theta_j + \frac{1}{2} D_{ijk} \theta_j \theta_k, \quad (10)$$

where $D_{ijk} = A_{ij,k}$. The first and second flexion are given in terms of derivatives of the two shear components (as per the notation in Bacon et al. 2006)

$$\mathcal{F} = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} \gamma_{1,1} - \gamma_{2,2} \\ \gamma_{2,1} + \gamma_{1,2} \end{pmatrix}. \quad (11)$$

The mass sheet degeneracy arises due to the fact that the observed quantities are not the shear and flexions themselves, but the reduced shear and the reduced first and second flexion. The latter two can be compactly written if we use a complex notation, i.e. $\mathcal{F} = \mathcal{F}_1 + i\mathcal{F}_2$,

$$G_1 = \frac{\mathcal{F} + g\mathcal{F}^*}{1 - \kappa}, \quad G_3 = \frac{\mathcal{G} + g\mathcal{F}}{1 - \kappa}, \quad (12)$$

where the asterisk denotes complex conjugation (Schneider & Er 2008). Alternatively, the reduced flexions can be defined as (Okura et al. 2008)

$$F = \frac{\mathcal{F}}{1 - \kappa}, \quad G = \frac{\mathcal{G}}{1 - \kappa}, \quad (13)$$

and we will use this definition in this paper. If the deflection potential is transformed for any constant λ as

$$\psi(\theta) \rightarrow \psi'(\theta) = \frac{1 - \lambda}{2} \theta^2 + \lambda \psi(\theta), \quad (14)$$

we have

$$\begin{aligned} \kappa \rightarrow \kappa' &= \lambda \kappa + (1 - \lambda), \quad \gamma \rightarrow \gamma' = \lambda \gamma, \\ \mathcal{F} \rightarrow \mathcal{F}' &= \lambda \mathcal{F}, \quad \mathcal{G} \rightarrow \mathcal{G}' = \lambda \mathcal{G}, \quad \mu \rightarrow \mu' = \frac{\mu}{\lambda^2}, \end{aligned} \quad (15)$$

but the reduced shear and reduced flexions are invariant under such a family of transformations,

$$g \rightarrow g' = g, \quad F \rightarrow F' = F, \quad G \rightarrow G' = G. \quad (16)$$

Therefore two different surface mass distributions that differ by λ cannot be distinguished by measuring only the image distortions. Note that in the limit of λ almost equal to unity, equation (15) amounts to adding a constant mass sheet κ_0 to κ .

In principle, the mass-sheet degeneracy could be lifted by requiring that $\kappa = 0$ at the boundary of the image. However, even for wide-field cameras this leads to a substantial underestimate of the cluster mass as cluster density profiles appear to be falling smoothly even out to large radius. For a cluster with virial mass $M_{\text{vir}} = 10^{15} M_{\odot}$ at redshift $z = 0.2$, we expect from N-body simulations that $\kappa \approx 0.005$ at $15'$ from the cluster center and with a $30' \times 30'$ camera field of view, we would underestimate the virial mass by $\sim 20\%$ if we simply set κ to zero at the boundary (Brađač et al. 2004). Therefore more sophisticated techniques are required.

3 PROPOSED METHOD: MOMENTS OF THE CONVERGENCE, SHEAR AND FLEXION FIELDS

Our goal is to derive a method that is independent of spectroscopic redshifts and broadly applicable to both critical and sub-critical lenses. To develop such a method, we first investigate n -th order moments of the shear, convergence, and flexions and study their transformation properties with a view to understanding the explicit dependence on λ . Schneider & Bartelmann (1997) (hereafter called SB97) show the equivalence of shear and convergence mass moments $M^{(n)}$ and multipole moments $Q^{(n)}$. We extend these equivalences to mass and multipole moments using flexion. Subsequently, we show that the mass-sheet degeneracy transformations destroy these equivalences, as they give rise to different surface terms. These permit us to calculate the mass-sheet degeneracy parameter λ .

Our method comes in two variants. One uses the reconstructed, unreduced shear or flexion fields. We can obtain these by multiplying the reduced quantities with $1 - \kappa_{\text{rec}}$, where κ_{rec} is the convergence of the reconstructed mass-sheet. Many techniques for obtaining κ_{rec} from shear

or flexion data have been successfully developed and Er et al. (2010) present a method that reconstructs the mass-sheet by combining the two. The second variant uses the reduced shear and flexion fields, thus it does not require the reconstructed convergence.

3.1 Mass and multipole moments of shear, flexion, and convergence

The detailed derivation of the different moments and the proof of their equivalences is shown in appendix A1. We do not explicitly denote (x_0, φ_0) or (x, φ) dependencies of $\kappa, \kappa_{\text{rec}}, \gamma, \mathcal{F}$, and \mathcal{G} to keep the notation compact. The resulting equivalent moments are:

$$M_{\kappa}^{(n)} = \int_0^{\infty} dx \, x^{n+1} w(x) \int_0^{2\pi} d\varphi \, \kappa, \quad (17)$$

$$M_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx \, x W(x) \int_0^{2\pi} d\varphi \, (1 - \kappa_{\text{rec}}) F_t, \quad (18)$$

$$M_{\gamma}^{(n)} = \int_0^{\infty} dx \, [2W(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi \, (1 - \kappa_{\text{rec}}) g_t, \quad (19)$$

$$M_{\mathcal{F}, \mathcal{G}}^{(n)} = - \int_0^{\infty} dx \, [V(x) - \frac{1}{2} x W(x)] \int_0^{2\pi} d\varphi \, [(1 - \kappa_{\text{rec}}) F_t + (1 - \kappa_{\text{rec}}) G_t]. \quad (20)$$

The formula for $M_{\mathcal{F}}^{(n)}$ in the case that $n = 0$ was also derived in Leonard et al. (2009), but with stronger assumptions on the weighting function. We used the definitions

$$x W(x) = \int_0^x dy \, y^{n+1} w(y), \quad (21)$$

$$V(x) = \int_0^x dy \, W(y), \quad (22)$$

and the reduced tangential and radial shear and flexions,

$$g_t = -[g_1(x, \varphi) \cos(2\varphi) + g_2(x, \varphi) \sin(2\varphi)], \quad (23)$$

$$g_r = -[g_2(x, \varphi) \cos(2\varphi) - g_1(x, \varphi) \sin(2\varphi)], \quad (24)$$

$$F_t = -[F_1(x, \varphi) \cos(\varphi) + F_2(x, \varphi) \sin(\varphi)], \quad (25)$$

$$F_r = -[F_2(x, \varphi) \cos(\varphi) - F_1(x, \varphi) \sin(\varphi)], \quad (26)$$

$$G_t = -[G_1(x, \varphi) \cos(3\varphi) + G_2(x, \varphi) \sin(3\varphi)], \quad (27)$$

$$G_r = -[G_2(x, \varphi) \cos(3\varphi) - G_1(x, \varphi) \sin(3\varphi)]. \quad (28)$$

We require that the boundary term $x W(x) \kappa(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. The derivation of M_{γ} demands in addition that the surface terms $x W(x) \gamma_t(x, \varphi)$ vanishes for the same limits and the derivation of $M_{\mathcal{F}, \mathcal{G}}$ demands in addition to the former two that $[2V(x) - x W(x)] \gamma_t(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. This can be achieved by using a suitable weighting function.

Analogously, we derived the equivalences between the

multipole moments

$$Q_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa, \quad (29)$$

$$Q_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) F_t, \quad (30)$$

$$Q_{\gamma}^{(n)} = \int_0^{\infty} dx [2W(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) g_t - i n \int_0^{\infty} dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) g_r, \quad (31)$$

$$Q_{\mathcal{F},G}^{(n)} = - \int_0^{\infty} dx [V(x) - \frac{1}{2} x W(x)] \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}) F_t + (1 - \kappa_{\text{rec}}) G_t] + i \frac{n}{2} \int_0^{\infty} dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}) F_r + (1 - \kappa_{\text{rec}}) G_r]. \quad (32)$$

Again we require that the term $xW(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. The derivation of Q_{γ} demands in addition that the surface term $xW(x)\gamma_t(x, \varphi)$ vanishes and the derivation of $Q_{\mathcal{F},G}$ demands in addition to the former two that $[2V(x) - xW(x)]\gamma_t(x, \varphi)$ and $V(x)\gamma_r(x, \varphi)$ vanish for $x \rightarrow 0, \infty$.

In some cases, it is useful to compute the moments only on a ring which excludes the central part of the field of view, for example if we want to exclude the strong lensing area. Therefore we extended the identities shown in this section to rings. The results are given in appendix A1.1.

3.2 Mass and multipole moments of reduced shear, reduced flexion, and K

The derivation of the moments and their equivalences is given in appendix B. The equivalent reduced moments are:

$$M_K^{(n)} = \int_R^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi K, \quad (33)$$

$$M_F^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi F_t, \quad (34)$$

$$M_g^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi (g_t F_t + g_r F_r) + \int_R^{\infty} dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi g_t, \quad (35)$$

$$M_{F,G}^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi (g_r F_r + \frac{1}{2}(F_t + G_t)) - \int_R^{\infty} dx V_R(x) \int_0^{2\pi} d\varphi (F_t + G_t - 2g_t F_t). \quad (36)$$

We defined analogous to Cain et al. (2011)

$$K = -\ln(1 - \kappa) \quad (37)$$

and in addition

$$xW_R(x) = \int_R^x dy y^{n+1} w(y), \quad (38)$$

$$V_R(x) = \int_R^x dy W_R(y). \quad (39)$$

We choose the lower integral limit for our ring, R , such that $\kappa < 1 \forall x \geq R, \forall \varphi$. Thus K is well-defined and finite. We require that $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. The M_g moment requires in addition that $xW_R(x)g_t(x, \varphi)$ vanishes for $x \rightarrow R, \infty$ and the $M_{F,G}$ moment demands that additionally to the two requirements also the surface term $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ vanishes for the same limits.

For the multipole moments, we have the following equivalent moments:

$$Q_K^{(n)} = \int_R^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} K, \quad (40)$$

$$Q_F^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} F_t, \quad (41)$$

$$Q_g^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t F_t + g_r F_r) + \int_R^{\infty} dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi e^{in\varphi} g_t - i n \int_R^{\infty} dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} g_r, \quad (42)$$

$$Q_{F,G}^{(n)} = \int_R^{\infty} dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t F_t + g_r F_r) - \int_R^{\infty} dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi e^{in\varphi} (\frac{1}{2}(F_t + G_t) - g_t F_t) + i n \int_R^{\infty} dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (\frac{1}{2}(F_r + G_r) - g_r F_r). \quad (43)$$

We have the same requirement on R as for the mass moments and we demand that the resulting surface term $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. For the Q_g moment we require additionally that $xW_R(x)g_t(x, \varphi)$ vanishes for the same limits. The $Q_{F,G}$ moment requires in addition to the two that the terms $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ and $V_R(x)g_r(x, \varphi)$ vanish for $x \rightarrow R$ and $x \rightarrow \infty$.

We note that the g moments require additional terms involving the flexion, which is not the case for the γ moments. Similarly, the F, G moments require shear information, which is not required for the \mathcal{F}, \mathcal{G} moments.

3.3 Mass moment transformations

3.3.1 Mass moments of shear, flexion, and convergence

The mass sheet degeneracy transformations in equation (15) destroy the mass moment equivalences shown in the previous section. They hold only for $\lambda = 1$. The n -th order mass

moment of the convergence for the transformed field κ' is

$$M_{\kappa'}^{(n)} = \int_0^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi (\lambda\kappa + (1-\lambda)) \quad (44)$$

and can be written explicitly in terms of the corresponding n -th order mass moments of κ ,

$$M_{\kappa'}^{(n)} = \lambda M_{\kappa}^{(n)} + 2\pi(1-\lambda) \int_0^\infty dx x^{n+1} w(x). \quad (45)$$

The integral on the right-hand side of the equation above is a Γ function for the choice of a Gaussian window function,

$$w(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad (46)$$

therefore we have

$$M_{\kappa'}^{(n)} = \lambda M_{\kappa}^{(n)} + 2^{\frac{n+1}{2}} \sqrt{\pi} (1-\lambda) \sigma^{n+2} \Gamma(1 + \frac{n}{2}). \quad (47)$$

Now let us examine the transformation properties for the mass moments computed using the shear field γ . In the limit that λ goes to unity, $M_{\kappa'}^{(n)}$ is equivalent to $M_{\gamma'}^{(n)}$. However, we show here that this is not the case in general, since the surface terms do not vanish identically. The moments for a transformed shear field $\gamma'_t = g_t (1 - \kappa'_{rec}) = \lambda g_t (1 - \kappa_{rec}) = \lambda \gamma_t$ are

$$\begin{aligned} M_{\gamma'}^{(n)} &= \int_0^\infty dx [2W(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi \gamma'_t \\ &= \lambda M_{\gamma}^{(n)}. \end{aligned} \quad (48)$$

Similarly, the mass moments in terms of flexion and convergence are equivalent in the limit $\lambda \rightarrow 1$, but they differ in general. For the flexion fields $\mathcal{F}'_t = F_t (1 - \kappa'_{rec}) = \lambda \mathcal{F}_t$ and $\mathcal{G}'_t = G_t (1 - \kappa'_{rec}) = \lambda \mathcal{G}_t$ we have

$$\begin{aligned} M_{\mathcal{F}'}^{(n)} &= \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi \mathcal{F}'_t \\ &= \lambda M_{\mathcal{F}}^{(n)}, \\ M_{\mathcal{F}', \mathcal{G}'}^{(n)} &= - \int_0^\infty dx [V(x) - \frac{1}{2} x W(x)] \int_0^{2\pi} d\varphi [\mathcal{F}'_t + \mathcal{G}'_t] \\ &= \lambda M_{\mathcal{F}, \mathcal{G}}^{(n)}. \end{aligned} \quad (49)$$

Therefore, by comparing the mass moments of a given order derived from the convergence field with that derived from the shear or flexion fields, λ can be estimated and hence the mass degeneracy can be lifted. A comparison of the shear and flexion mass moments cannot be used to infer λ , because they are equivalent for every λ .

As shown in appendix A1, the requirements on the weighting function for the use of $M_{\mathcal{F}, \mathcal{G}}^{(n)}$ are stricter than those for $M_{\mathcal{F}}^{(n)}$. The application of the former demands in the case of a simple Gaussian weighting function that the tangential shear decreases faster than $1/\log(x)$ for $x \rightarrow \infty$ to ensure that equation (A31) holds, while the use of the latter does not require such constraints. Shear typically decreases as $1/x$ and both identities are therefore usable, but it can be advantageous to design a better suited weighting function.

3.3.2 Mass moments of reduced shear, reduced flexion, and K

The mass sheet degeneracy transformations also break the equivalences of the mass moments of the reduced quantities. We have

$$\begin{aligned} K' &= -\ln(1 - (\lambda\kappa + 1 - \lambda)) \\ &= -\ln(\lambda) - \ln(1 - \kappa) \end{aligned} \quad (51)$$

and therefore the corresponding mass moment transforms as

$$M_{K'}^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi [-\ln(\lambda) - \ln(1 - \kappa)] \quad (52)$$

$$= M_K^{(n)} - 2\pi \ln(\lambda) \int_R^\infty dx x^{n+1} w(x). \quad (53)$$

For a Gaussian weighting function, this leads to

$$M_{K'}^{(n)} = M_K^{(n)} - 2^{\frac{n+1}{2}} \sqrt{\pi} \ln(\lambda) \sigma^{n+2} \Gamma(1 + \frac{n}{2}, \frac{R^2}{2\sigma^2}), \quad (54)$$

where $\Gamma(a, x)$ is the incomplete upper gamma function,

$$\Gamma(a, x) = \int_x^\infty dt t^{a-1} e^{-t}. \quad (55)$$

The remaining mass moments depend only on reduced quantities which do not change under this type of transformation, thus we have

$$M_{g'}^{(n)} = M_g^{(n)}, \quad (56)$$

$$M_{F'}^{(n)} = M_F^{(n)}, \quad (57)$$

$$M_{F', G'}^{(n)} = M_{F, G}^{(n)}. \quad (58)$$

As we showed in the previous section, for $\lambda \rightarrow 1$ all moments of a given order are equal and we can therefore compare one of the latter three to the K mass moment to determine λ .

As seen in appendix B, the requirements for the use of $M_{F, G}^{(n)}$ are again stricter than those for $M_F^{(n)}$. For a Gaussian weighting function, the reduced tangential shear has to decrease faster than $1/\log(x)$ for $x \rightarrow \infty$, which is typically the case. Designing a faster declining weight can avoid the requirement on the reduced shear.

3.4 Multipole moment transformations

3.4.1 Multipole moments of shear, flexion, and convergence

We show that the multipole moments are equivalent under the mass sheet transformations for all λ . Therefore they cannot be used to determine its value. This is due to the additional exponential factor $e^{in\varphi}$. For the following calculation, we will assume $n \geq 1$, as the multipole moments are identical to the mass moments for $n = 0$. We have

$$\begin{aligned} Q_{\kappa'}^{(n)} &= \int_0^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\lambda\kappa + (1-\lambda)] \\ &= \lambda Q_{\kappa}^{(n)}. \end{aligned} \quad (59)$$

The shear and flexion multipole moments transform like their mass moments,

$$Q_{\gamma'}^{(n)} = \lambda Q_{\gamma}^{(n)}, \quad (60)$$

$$Q_{\mathcal{F}'}^{(n)} = \lambda Q_{\mathcal{F}}^{(n)}, \quad (61)$$

$$Q_{\mathcal{F}',\mathcal{G}'}^{(n)} = \lambda Q_{\mathcal{F},\mathcal{G}}^{(n)}. \quad (62)$$

3.4.2 Multipole moments of reduced shear, reduced flexion, and K

The multipole moments obtained from the reduced fields are also equivalent for every choice of λ and thus they cannot be used to break the mass sheet degeneracy. As before, we restrict ourselves to the multipole moments with $n \geq 1$, because for $n = 0$ they are identical to the mass moments. The K multipole moment transforms as

$$\begin{aligned} Q_{K'}^{(n)} &= \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \\ &\quad [-\ln(\lambda) - \ln(1 - \kappa)] \\ &= Q_K^{(n)}. \end{aligned} \quad (63)$$

Like the mass moments, the remaining multipole moments do not change under a transformation of the mass-sheet:

$$Q_{g'}^{(n)} = Q_g^{(n)}, \quad (64)$$

$$Q_{F'}^{(n)} = Q_F^{(n)}, \quad (65)$$

$$Q_{F',G'}^{(n)} = Q_{F,G}^{(n)}. \quad (66)$$

4 APPLICATION TO A SIS MODEL

We apply our theoretical framework to a singular isothermal sphere (SIS) model to illustrate the method. The projected surface mass density is given by (Bartelmann & Schneider 2001):

$$\Sigma_{\text{SIS}}(\xi) = \frac{\sigma_\nu^2}{2G\xi}, \quad (67)$$

where ξ is the separation from the lens center in the projected lens plane and σ_ν is the line-of-sight velocity dispersion of the particles. The convergence is given by $\kappa(\xi) = \Sigma(\xi)/\Sigma_{\text{crit}}$ and defining the angular distance $\theta = \xi/D_{\text{ol}}$ and the Einstein deflection angle

$$\theta_E = 4\pi \left(\frac{\sigma_\nu}{c} \right)^2 \frac{D_{\text{ls}}}{D_{\text{os}}} \quad (68)$$

leads to

$$\kappa_{\text{SIS}}(\theta) = \frac{\theta_E}{2\theta}. \quad (69)$$

The shear for a SIS at a vectorial angular separation θ is (Bartelmann & Schneider 2001)

$$\gamma_{\text{SIS}}(\theta) = -\frac{\theta_E}{2\theta} e^{2i\varphi} \quad (70)$$

and for flexion, we have (Bacon et al. 2006)

$$\mathcal{F}_{\text{SIS}}(\theta) = -\frac{\theta_E}{2\theta^2} e^{i\varphi}, \quad \mathcal{G}_{\text{SIS}}(\theta) = \frac{3\theta_E}{2\theta^2} e^{3i\varphi}, \quad (71)$$

where we used the complex shear notation $\gamma = \gamma_1 + i\gamma_2$ and the analogous notation for flexion. Therefore the tangential and radial components are

$$\gamma_{t,\text{SIS}}(\theta) = \frac{\theta_E}{2\theta}, \quad \gamma_{r,\text{SIS}}(\theta) = 0, \quad (72)$$

$$\mathcal{F}_{t,\text{SIS}}(\theta) = \frac{\theta_E}{2\theta^2}, \quad \mathcal{F}_{r,\text{SIS}}(\theta) = 0, \quad (73)$$

$$\mathcal{G}_{t,\text{SIS}}(\theta) = -\frac{3\theta_E}{2\theta^2}, \quad \mathcal{G}_{r,\text{SIS}}(\theta) = 0. \quad (74)$$

Measuring higher order mass moments tends to be a noisy affair, therefore we examine for simplicity the lowest order mass moment $M^{(0)}$. Since the 0-th order mass and multipole moments are identical, we will also restrict our treatment to the mass moment. We investigate the non-reduced moment method, as these moments give short analytic expressions and the reduced moment technique is exactly analogous. Writing the tangential components in terms of θ and performing the integrals using the Gaussian weighting function, we have for the SIS

$$M_{\kappa,\text{SIS}}^{(0)} = M_{\gamma,\text{SIS}}^{(0)} = M_{\mathcal{F},\text{SIS}}^{(0)} = M_{\mathcal{F},\mathcal{G},\text{SIS}}^{(0)} = \frac{\pi \sigma \theta_E}{2}. \quad (75)$$

For the transformed moments, we have

$$M_{\kappa'}^{(0)} = \lambda \frac{\pi \sigma \theta_E}{2} + \sqrt{2\pi}(1 - \lambda)\sigma^2, \quad (76)$$

$$M_{\gamma',\text{SIS}}^{(0)} = M_{\mathcal{F}',\text{SIS}}^{(0)} = M_{\mathcal{F}',\mathcal{G}',\text{SIS}}^{(0)} = \lambda \frac{\pi \sigma \theta_E}{2}. \quad (77)$$

Therefore we can estimate the value of $(1 - \lambda)$ by evaluating $(M_{\kappa'}^{(0)} - M_{\kappa,\text{SIS}}^{(0)})/\sqrt{2\pi}\sigma^2$, where x represents a shear or flexion mass moment.

5 DETECTABILITY

In this section, we estimate the reliability of λ detections from the mass moment measurements *without assuming a halo model*. For this purpose, we compare the moments with the moment standard deviations in the absence of lensing. The latter are calculated in appendix C. We assume that the convergence information is derived from number counts exploiting the magnification bias and that the unlensed number counts follow a power law $n_0 \propto S^{-\beta}$ with flux limit S and $\beta = 0.5$ (e.g., Schneider et al. 2000). The results are

$$\sigma_{M,\kappa} = \sqrt{\frac{2\pi}{n_\kappa}} \left(\int_0^\infty dx x^{2n+1} w(x)^2 \right)^{\frac{1}{2}}, \quad (78)$$

$$\sigma_{M,\gamma} = \sqrt{\frac{2\pi}{n_\gamma}} \sigma_\gamma \left(\int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \right)^{\frac{1}{2}}, \quad (79)$$

$$\sigma_{M,\mathcal{F}} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}}} \sigma_{\mathcal{F}} \left(\int_0^\infty dx x W(x)^2 \right)^{\frac{1}{2}}, \quad (80)$$

$$\begin{aligned} \sigma_{M,\mathcal{F},\mathcal{G}} &= \sqrt{2\pi} \sqrt{\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}}} \\ &\quad \left(\int_0^\infty dx \frac{1}{x} [V(x) - \frac{1}{2}xW(x)]^2 \right)^{\frac{1}{2}} \end{aligned} \quad (81)$$

for the non-reduced moments and

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 A^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}, \quad (82)$$

$$\sigma_{M,g} = \left(\frac{4\pi}{n_\mathcal{F}} \sigma_\mathcal{F}^2 \sigma_\gamma^2 \left(\int_R^\infty dx x W_R(x)^2 \right) + \frac{2\pi}{n_\gamma} \sigma_\gamma^2 \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1} w(x))^2 \right)^{\frac{1}{2}}, \quad (83)$$

$$\sigma_{M,F} = \sqrt{\frac{2\pi}{n_\mathcal{F}}} \sigma_\mathcal{F} \left(\int_R^\infty dx x W_R(x)^2 \right)^{\frac{1}{2}}, \quad (84)$$

$$\begin{aligned} \sigma_{M,F,G} = & \sqrt{2\pi} \left(\left(\frac{\sigma_\mathcal{F}^2 \sigma_\mathcal{F}^2}{n_\mathcal{F}} + \frac{\sigma_\mathcal{F}^2}{4n_\mathcal{F}} + \frac{\sigma_\mathcal{G}^2}{4n_\mathcal{G}} \right) \int_R^\infty dx x W_R(x)^2 \right. \\ & \left. - \left(\frac{\sigma_\mathcal{F}^2}{n_\mathcal{F}} + \frac{\sigma_\mathcal{G}^2}{n_\mathcal{G}} \right) \int_R^\infty dx W_R(x) V_R(x) \right. \\ & \left. + \left(\frac{\sigma_\mathcal{F}^2}{n_\mathcal{F}} + \frac{\sigma_\mathcal{G}^2}{n_\mathcal{G}} + 4 \frac{\sigma_\mathcal{F}^2 \sigma_\mathcal{F}^2}{n_\mathcal{F}} \right) \int_R^\infty dx \frac{1}{x} V_R(x)^2 \right)^{\frac{1}{2}} \quad (85) \end{aligned}$$

for the reduced moments. Here A denotes our total field of view area. Using the 2 techniques described in this paper, we can measure $1-\lambda$ or $\ln(\lambda)$ and their respective detectabilities are given by:

$$\left(\frac{S}{N} \right)_{1-\lambda}^{(n)} = \left| \frac{M_{\kappa'}^{(n)} - M_{x'}^{(n)}}{(\sigma_{M^{(n)},\kappa'}^2 + \sigma_{M^{(n)},x'}^2)^{\frac{1}{2}}} \right|, \quad (86)$$

$$\left(\frac{S}{N} \right)_{\ln(\lambda)}^{(n)} = \left| \frac{M_{K'}^{(n)} - M_{X'}^{(n)}}{(\sigma_{M^{(n)},K'}^2 + \sigma_{M^{(n)},X'}^2)^{\frac{1}{2}}} \right|, \quad (87)$$

where x and X represent the non-reduced and reduced shear and flexion moments used to break the degeneracy. We calculate the detectabilities for the lowest order mass moment combinations. We assume a λ of 1.2 and that we can measure the shear in 100 sources/arcmin², as in the *Hubble Frontier Fields*. We further choose $n_\kappa = 300/\text{arcmin}^2$, $\sigma_\gamma = 0.26$ (Leauthaud et al. 2007), and $n_\mathcal{F} = n_\mathcal{G} = 25/\text{arcmin}^2$. We have $\sigma_{F a_{\text{gal}}} \approx \sigma_{\mathcal{F} a_{\text{gal}}}$ and $\sigma_{G a_{\text{gal}}} \approx \sigma_{\mathcal{G} a_{\text{gal}}}$, where a_{gal} is the semi-major of the lensed source, and typical values for the intrinsic flexion of 0.03 and 0.04 respectively (Goldberg & Leonard 2007; Goldberg & Bacon 2005). As flexion can more reliably be measured for larger sources, we assume a lower source density and $a_{\text{gal}} = 0.2''$. This leads to $\sigma_\mathcal{F} = 0.15$ 1/arcsec and $\sigma_\mathcal{G} = 0.2$ 1/arcsec. For the galaxy lens, we assume an Einstein radius of $0.2''$ and for the cluster $20''$, but note that our results do not require any assumptions on the halo shape, as the difference between the mass moments solely depends on λ and the weighting function. We choose the lower integration boundary R to be the Einstein radius and for the noise integrals involving intrinsic flexion, we choose an upper integration boundary of 3 arcmin for the galaxy lens and of 30 arcmin for the cluster instead

of infinity. We can do this because flexion drops off very quickly, typically as $1/x^2$, and therefore we can assume that the flexion and thus its noise is 0 outside of a certain, generously chosen area. E.g. for an SIS, we have that the \mathcal{F} and \mathcal{G} signals drop to less than 10^{-5} 1/arcsec at these distances. If we do not make this simplification, we would integrate the noise out to infinity, as the W weighting function does not drop off quickly enough for a Gaussian w function, even though there is obviously no signal in this region. In addition, we assume that we have an image area of 11 arcmin², which corresponds to the field of view of the *Hubble Space Telescope Advanced Camera for Surveys*.

For the stacking calculations, we assume that we stack 1000 galaxy-scale lenses to obtain the convergence map of an averaged galaxy lens. In this case, we can compute the λ parameter and lift the mass sheet degeneracy for the stacked, i.e. averaged, lensing galaxy.

Table 1 shows the detectability of $1-\lambda$ and $\ln(\lambda)$ for a Gaussian weighting function w and several choices of σ . A larger σ leads to a wider weighting function that encompasses more of the lensing information and thus results in a larger detectability. If the convergence and nonreduced shear moments are combined, we clearly detect λ and we can break the mass sheet degeneracy for stacked galaxy lenses and cluster lenses. Similarly, we can estimate λ for stacking if we combine K and reduced shear or if we use the first flexion moments, even though the latter are much noisier.

However, in several cases the detectability is too low to reliably measure λ . The reason is that these moment combinations involve the flexion and thus the weighting function $W(x) = 1/x \int_0^x dy y^{n+1} w(y)$ or

$W_R(x) = 1/x \int_R^x dy y^{n+1} w(y)$, respectively. For a Gaussian weight $w(x)$ and $n = 0$, the integral rapidly approaches a constant value. Thus $W(x)$ and $W_R(x)$ drop off very slowly as $1/x$, while the flexion drops off as $1/x^2$. As a result, a large area which contains no signal is highly weighted. This results in a low detectability. If a faster declining W function is designed, e.g. by using a $w(x)$ which becomes negative at a certain distance from the lens center, the detectability of these moments can be substantially improved. This is in particular true for the mass moments using first and second flexion, as these require in addition the $V(x)$ and $V_R(x)$ weights, which even increase with distance. The mass moments remain finite, as the flexion drops off faster than the weight increases, but a quickly decreasing weight would increase the detectability significantly.

A second reason for using fast declining W and V functions is the following: In our calculation, we could safely neglect very small flexion signals without affecting the resulting mass moment difference, because we assumed a value for λ . However, depending on σ and the mass moment in question, these very small flexion values can be amplified by the slow drop off of the weighting function and neglecting them can lead to a potentially substantial error in the mass moment estimate.

Table 1. Detectability of $1 - \lambda$ (upper part) and $\ln(\lambda)$ (lower part) for different moment combinations. The results are shown for galaxy lenses, 1000 stacked galaxy lenses, and cluster lenses for different σ values of the Gaussian $w(x)$ weighting function. The moment combinations listed after the first one involve the W or V weighting function types, which drop off very slowly. Thus their detectability can be substantially improved by designing a faster declining weighting function (see text).

	Galaxy lensing			Stacking			Cluster lensing		
σ	2''	5''	10''	2''	5''	10''	20''	50''	100''
$\kappa + \gamma$	0.4	0.9	1.9	11.8	29.5	59.0	3.7	9.3	18.7
$\kappa + \mathcal{F}$	0.1	0.2	0.2	4.1	4.9	5.6	0.1	0.2	0.2
$\kappa + \mathcal{F}, \mathcal{G}$	0.04	0.05	0.08	1.2	1.7	2.6	0.04	0.05	0.08
$K + g$	0.2	0.4	0.4	7.6	11.1	13.4	0.3	0.4	0.4
$K + F$	0.1	0.1	0.2	3.8	4.5	5.1	0.1	0.1	0.2
$K + F, G$	0.03	0.05	0.07	1.0	1.5	2.1	0.04	0.05	0.07

Several optimized aperture mass weighting functions have been proposed (e.g., [Leonard et al. 2009](#); [Schneider & Bartelmann 1997](#)), but they typically assume that the weight w is compensated, i.e. $\int_0^\infty xw(x)dx = 0$. We did not impose this condition on the weighting function in our derivations, and indeed a compensated w cannot be used to break the mass sheet degeneracy using the non-reduced moments of lowest order, which are expected to have the minimal noise, as the additional term in equation (45) would vanish.

[Maturi et al. \(2005\)](#) propose a non-compensated, optimized weighting function for shear that takes into account the shape of the signal and the noise power spectrum. This approach is promising, but it is complicated by the fact that for our technique the weights w , W , and V are related to each other. Thus it is not possible to optimize the weighting functions independently. In fact, we must optimize at least two inter-dependent weighting functions for very different signal and noise shapes simultaneously. This optimization is outside the scope of this paper and is therefore left for a future publication.

Lensing measurements in observational data give discrete sets rather than smooth functions of observables. An application of the aperture mass technique to observational data can therefore be achieved by either binning and approximating the integral with a Riemann sum or fitting the results to a smooth function.

6 CONVERGENCE, SHEAR, AND FLEXION SIGNAL-TO-NOISE

We compare the signal-to-noise ratios for SIS and Navarro-Frenk-White (NFW) profiles ([Navarro et al. 1996, 1997](#)). We show that the combination of shear and flexion has a higher signal-to-noise ratio than the number counts method for both galaxy and cluster lensing. Therefore including number counts in the mass sheet reconstruction does not improve the lens model. Thus we can use this information to break the mass sheet degeneracy without sacrificing the accuracy of the model.

We can restrict ourselves to the non-reduced shear and flexion, because the extra $1/(1 - \kappa)$ term in the reduced quantities would only boost their signal-to-noise. The convergence, shear, and flexion formulas for a NFW profile are given in appendix D. In addition, we only investigate the intrinsic noise of each method and neglect possible additional measurement errors.

We can make flexion dimensionless by multiplying it with the semi-major of the observed galaxy a_{gal} ([Goldberg & Bacon 2005](#)). The signal-to-noise ratios are

$$\left(\frac{S}{N}\right)_{Fa_{\text{gal}}} = \frac{|F| a_{\text{gal}} \sqrt{N_F}}{\sigma_{Fa_{\text{gal}}}}, \quad (88)$$

$$\left(\frac{S}{N}\right)_{Ga_{\text{gal}}} = \frac{|G| a_{\text{gal}} \sqrt{N_G}}{\sigma_{Ga_{\text{gal}}}}, \quad (89)$$

where $\sigma_{Fa_{\text{gal}}}$ and $\sigma_{Ga_{\text{gal}}}$ are the dispersions for dimensionless reduced flexion and we average over N_F and N_G sources. We can infer the signal-to-noise of $\mathcal{F}a_{\text{gal}}$ and $\mathcal{G}a_{\text{gal}}$ by replacing the reduced quantities with these expressions. We have again $\sigma_{Fa_{\text{gal}}} \approx \sigma_{\mathcal{F}a_{\text{gal}}}$ and $\sigma_{Ga_{\text{gal}}} \approx \sigma_{\mathcal{G}a_{\text{gal}}}$ with typical values of 0.03 and 0.04 respectively ([Goldberg & Leonard 2007](#); [Goldberg & Bacon 2005](#)). The signal-to-noise of the shear and number counts is ([Schneider et al. 2000](#))

$$\left(\frac{S}{N}\right)_g = \frac{|g| \sqrt{N_g}}{\sigma_\epsilon}, \quad (90)$$

$$\left(\frac{S}{N}\right)_\kappa = |\mu^{\beta-1} - 1| \sqrt{N_\kappa} \approx 2\kappa |1 - \beta| \sqrt{N_\kappa}, \quad (91)$$

where we have used the first order weak lensing expansion. We average our number counts over N_κ sources and β is the number counts exponent, see appendix C. As we investigate the non-reduced quantities, we replace g with γ . We choose $\sigma_\epsilon = 0.26$ ([Leauthaud et al. 2007](#)) and the remaining parameters analogously to [Schneider et al. \(2000\)](#), $\beta = 0.5$ and $N_\kappa = 3.5 N_g$. The latter is a slightly optimistic estimate for N_κ , since for real observations, we will lose several sources due to color cuts. A lower estimate would decrease the signal-to-noise of the number counts method, so we will keep this value. Using the expressions for the SIS, we have

$$\frac{(S/N)_\gamma}{(S/N)_\kappa} = 2.1, \quad (92)$$

$$\frac{(S/N)_{\mathcal{F}a_{\text{gal}}}}{(S/N)_\kappa} = \frac{a_{\text{gal}}}{0.03 \theta} \sqrt{\frac{N_F}{N_\kappa}}, \quad (93)$$

$$\frac{(S/N)_{\mathcal{G}a_{\text{gal}}}}{(S/N)_\kappa} = \frac{3 a_{\text{gal}}}{0.04 \theta} \sqrt{\frac{N_G}{N_\kappa}}. \quad (94)$$

The radial dependence of the ratios is shown in figure 1 for $a_{\text{gal}} = 0.2''$ and $N_F = N_G$ values of N_g , $0.5 N_g$, and $0.1 N_g$. Note that the signal-to-noise ratios do not depend on the Einstein radius θ_E , which is canceled in the calculation. However, the distance of the weak lensing regime from the halo center does depend on the Einstein radius. Therefore we show κ for Einstein radii of $0.2''$, $10''$, and $20''$. For larger Einstein radii, shear by itself has always the best signal-to-noise in the weak lensing regime. For smaller θ_E , shear and flexion together give the best results. Note that even in the case of large Einstein radii, the flexion information is not redundant. It is much more sensitive to small scale structures than shear and can therefore be

used to reconstruct them (Leonard et al. 2009; Bacon et al. 2010).

The analytic expressions for the NFW profile are complex and therefore it is more instructing to plot the signal-to-noise ratios. These are independent of ρ_{crit} , δ_c , and Σ_{crit} . We show the radial dependence for three angular scale radii $\theta_s = r_s/D_{\text{ol}}$ in figure 2. The first plot corresponds to a galaxy at redshift $z = 0.35$ with $r_s = 16 h^{-1}$ kpc and sources at redshift $z = 0.6$, which agrees e.g. with the observations in Hoekstra et al. (2004). The second and third correspond to clusters with $r_s = 250 h^{-1}$ kpc and $r_s = 70 h^{-1}$ kpc at the same redshift, comparable to the NFW models for one component of Abell 370 and for the cluster MS 2137 in Shu et al. (2008). For the first cluster, we only model one component to obtain a NFW profile with intermediate Einstein radius for our comparison. Assuming a standard, flat Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$, we calculate the angular diameter distances using the Wright (2006) cosmology calculator and astropy (Astropy Collaboration et al. 2013) and obtain the corresponding angular scale radii $4.6''$, $72''$, and $20''$. For the galaxy, we further set $\delta_c = 2.4 \cdot 10^4$ and for the two clusters $\kappa_s = \rho_{\text{crit}} \delta_c r_s / \Sigma_{\text{crit}} = 0.16$ and 0.66 , in agreement with the values cited in the respective publications. The size of the semi-major and the number of measured sources are the same as for the SIS. The effect of a different choice of these parameters can be visualized by a shift of the corresponding graphs, because the scale in the figures is logarithmic.

Figure 2 shows that shear by itself has always the best signal-to-noise ratio in the weak lensing regime of the 2 clusters. For the galaxy, a combination of flexion and shear provides the best results, as the flexion dominates at small separations from the center. This illustrates also the value of flexion for the reconstruction of substructures.

7 DISCUSSION AND CONCLUSIONS

We have derived mass and multipole moment formulas in terms of flexion and showed that they are equivalent to the mass and multipole moments of shear and convergence. Furthermore, we have showed the equivalence of mass and multipole moments in terms of reduced shear, reduced flexion, and K , which is a quantity derived from the convergence. We investigated the moment behavior with respect to the mass sheet degeneracy transformation and found that the equivalences are broken for the mass moments, but preserved for the multipole moments. The resulting surface terms can be used to break the mass sheet degeneracy.

We demonstrated the new theoretical framework by applying it to a SIS mass model. In addition, we investigated the detectability of the mass sheet parameter λ without assuming a halo model. We found that we can break the mass sheet degeneracy for stacked galaxy-galaxy and cluster lensing. Combinations of the shear and convergence moments have a much higher detectability than combinations using the flexions. This is due to their weighting function, which drops off too slowly. A fine-tuned weight

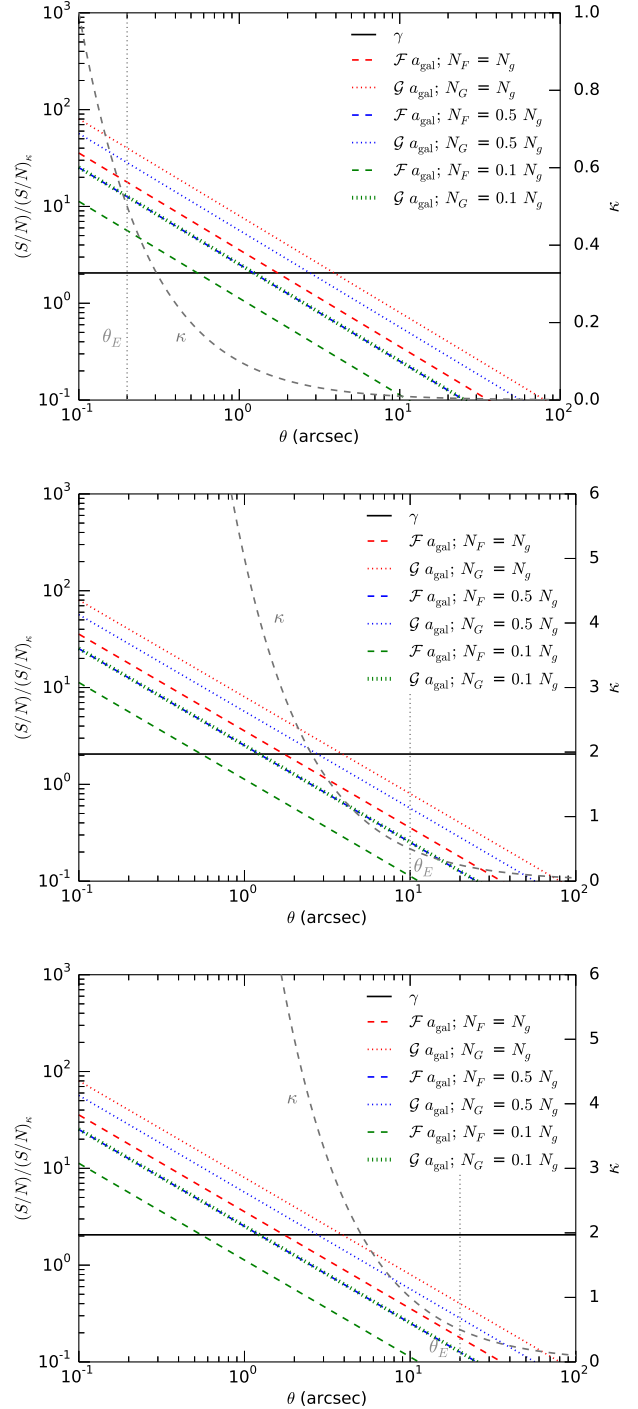


Figure 1. Comparison of the signal-to-noise ratios of γ , $\mathcal{F} a_{\text{gal}}$, and $\mathcal{G} a_{\text{gal}}$ normalized to $(S/N)_\kappa$, the signal-to-noise of number counts, for the singular isothermal sphere. The Einstein radii are $0.2''$ (top), $10''$ (middle), and $20''$ (bottom). We show the flexion results for source densities of 100%, 50%, and 10% of the shear source density and plot the convergence value as a reference.

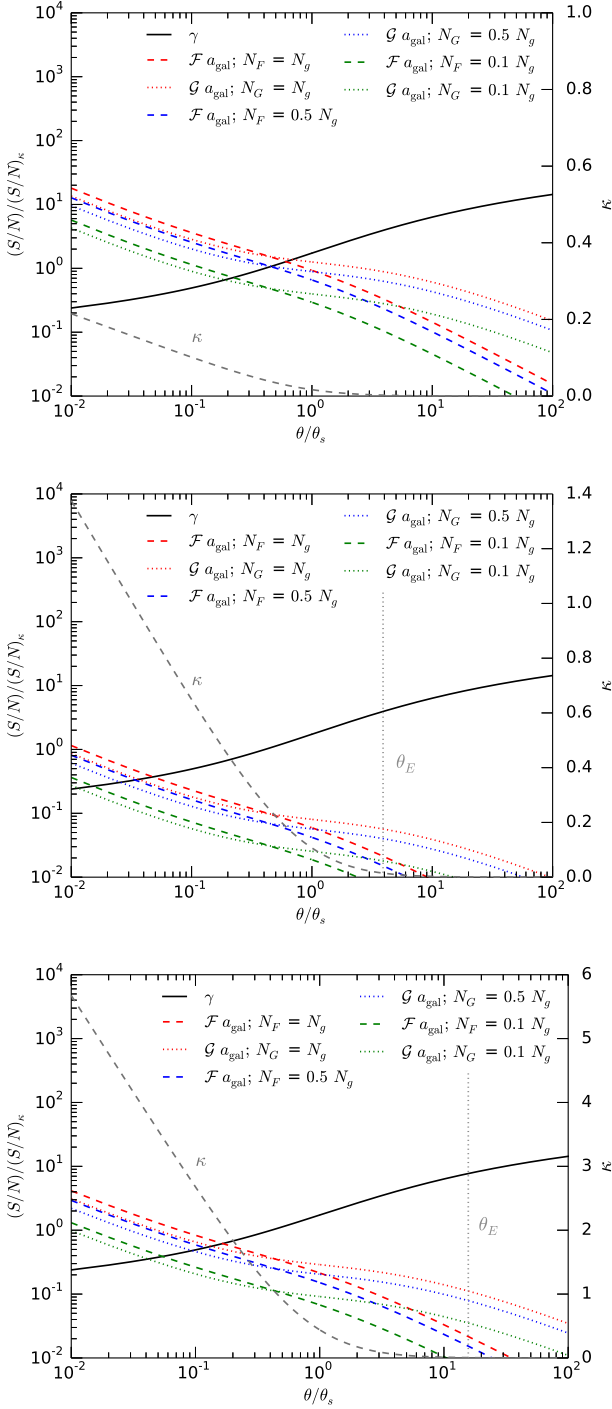


Figure 2. Comparison of (S/N) ratios of γ , $\mathcal{F} a_{\text{gal}}$, and $\mathcal{G} a_{\text{gal}}$ normalized to $(S/N)_\kappa$, the signal-to-noise of number counts, for NFW profiles with $\theta_s = 4.6''$ (top), $72''$ (middle), and $20''$ (bottom). The flexion is shown for source densities of 100%, 50%, and 10% of the shear source density. The convergence and the Einstein radius are plotted as a reference. The Einstein radius of the galaxy is too small to be shown.

will significantly improve their performance.

Finally, we investigated the signal-to-noise ratios of shear, flexion, and convergence information from number counts for SIS and NFW halos in the weak lensing regime. It was assumed that the noise in γ , \mathcal{F} , and \mathcal{G} is dominated by the intrinsic signal of the lensed sources and that the number counts are dominated by Poisson noise. The estimates do not include measurement errors e.g. due to pixel noise. This will have an additional effect, in particular on flexion and if sources with insufficient signal-to-noise are used (Rowe et al. 2013). With high quality data, e.g. from the *Hubble Frontier Fields*, the reliability of the flexion measurements should improve. Under these assumptions, we demonstrated that the combination of shear and flexion always gives better results than other combinations using the number counts. Using an estimate of shear and number count signal-to-noise ratios, e.g. Schneider et al. (2000) also found that shear is the superior method. However, Van Waerbeke et al. (2010) and Hildebrandt et al. (2013) caution that we can still count sources at redshifts for which shear measurements are no longer feasible, thus reducing the advantage of shear. We find that the convergence information can thus be used to break the mass sheet degeneracy without a loss in the quality of the lens model.

ACKNOWLEDGEMENTS

MR thanks the Department of Astronomy at Yale University for hosting him and acknowledges support by the German National Academic Foundation. He thanks Thibault Kuntzer, Austen Groener, and Justin Bird for fruitful discussions and Alexandre Refregier for his support. MR and PN acknowledge useful discussions with David Goldberg and Richard Massey on flexions. MR and JPK acknowledge support from the ERC advanced grant LIDA. PN gratefully acknowledges support from an NSF theory program via the grant AST-1044455 and a theory grant from the Space Telescope Science Institute HST-AR-12144.01-A.

REFERENCES

- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Bacon D. J., Goldberg D. M., Rowe B. T. P., Taylor A. N., 2006, *MNRAS*, **365**, 414
- Bacon D. J., Amara A., Read J. I., 2010, *MNRAS*, **409**, 389
- Bartelmann M., 1996, *A&A*, **313**, 697
- Bartelmann M., Narayan R., 1995, *ApJ*, **451**, 60
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, **340**, 291
- Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, *ApJ*, **464**, L115
- Bradač M., Lombardi M., Schneider P., 2004, *A&A*, **424**, 13
- Broadhurst T. J., Taylor A. N., Peacock J. A., 1995, *ApJ*, **438**, 49
- Cain B., Schechter P. L., Bautz M. W., 2011, *ApJ*, **736**, 43
- Dye S., et al., 2002, *A&A*, **386**, 12
- Er X., Li G., Schneider P., 2010, preprint ([arXiv:1008.3088](https://arxiv.org/abs/1008.3088))
- Fort B., Mellier Y., Dantel-Fort M., 1997, *A&A*, **321**, 353
- Goldberg D. M., Bacon D. J., 2005, *ApJ*, **619**, 741
- Goldberg D. M., Leonard A., 2007, *ApJ*, **660**, 1003
- Goldberg D. M., Natarajan P., 2002, *ApJ*, **564**, 65

- Goodman L. A., 1960, *Journal of the American Statistical Association*, 55, 708
- Hildebrandt H., et al., 2013, *MNRAS*, 429, 3230
- Hoekstra H., Yee H. K. C., Gladders M. D., 2004, *ApJ*, 606, 67
- Irwin J., Shmakova M., 2006, *ApJ*, 645, 17
- Kaiser N., 1995, *ApJ*, 439, L1
- Kaiser N., Squires G., 1993, *ApJ*, 404, 441
- Kneib J.-P., Natarajan P., 2011, *A&ARv*, 19, 47
- Leauthaud A., et al., 2007, *ApJS*, 172, 219
- Leonard A., Goldberg D. M., Haaga J. L., Massey R., 2007, *ApJ*, 666, 51
- Leonard A., King L. J., Wilkins S. M., 2009, *MNRAS*, 395, 1438
- Maturi M., Meneghetti M., Bartelmann M., Dolag K., Moscardini L., 2005, *A&A*, 442, 851
- Meylan G., Jetzer P., North P., eds, 2006, *Gravitational Lensing: Strong, Weak and Micro* ([arXiv:astro-ph/0407232](https://arxiv.org/abs/astro-ph/0407232)), <http://adsabs.harvard.edu/abs/2006glsw.conf.....M>
- Natarajan P., Kneib J.-P., Smail I., Ellis R. S., 1998, *ApJ*, 499, 600
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Okura Y., Umetsu K., Futamase T., 2007, *ApJ*, 660, 995
- Okura Y., Umetsu K., Futamase T., 2008, *ApJ*, 680, 1
- Rexroth M., 2015, submitted to *IAU Proceedings Astronomy in Focus*
- Rowe B., Bacon D., Massey R., Heymans C., Häußler B., Taylor A., Rhodes J., Mellier Y., 2013, *MNRAS*, 435, 822
- Schneider P., Bartelmann M., 1997, *MNRAS*, 286, 696
- Schneider P., Er X., 2008, *A&A*, 485, 363
- Schneider P., Seitz C., 1995, *A&A*, 294, 411
- Schneider P., King L., Erben T., 2000, *A&A*, 353, 41
- Shu C., Zhou B., Bartelmann M., Comerford J. M., Huang J.-S., Mellier Y., 2008, *ApJ*, 685, 70
- Taylor A. N., Dye S., Broadhurst T. J., Benitez N., van Kampen E., 1998, *ApJ*, 501, 539
- Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, *ApJ*, 729, 127
- Van Waerbeke L., Hildebrandt H., Ford J., Milkeraitis M., 2010, *ApJ*, 723, L13
- Wright E. L., 2006, *PASP*, 118, 1711
- Wright C. O., Brainerd T. G., 2000, *ApJ*, 534, 34

APPENDIX A: DERIVATION OF THE MASS AND MULTIPOLE MOMENT EQUIVALENCES FOR CONVERGENCE, SHEAR, AND FLEXION

A1 Moments on the full field of view

We compute the general n -th order moments within an aperture defined by the weighting function $w(x)$. This allows the application to finite field observational data. The moments of the mass distribution in terms of κ are defined as (SB97)

$$M_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \kappa(x, \varphi, x_0, \varphi_0) \quad (\text{A1})$$

and the corresponding multipole moments are defined as

$$Q_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa(x, \varphi, x_0, \varphi_0). \quad (\text{A2})$$

From now on we will no longer write the explicit dependence on x_0 and φ_0 to simplify the notation. As in (SB97), we can integrate equation (A1) by parts to obtain

$$M_{\kappa}^{(n)} = - \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi \frac{\partial \kappa(x, \varphi)}{\partial x}, \quad (\text{A3})$$

where

$$x W(x) = \int_0^x dy y^{n+1} w(y) \quad (\text{A4})$$

and we require that the boundary term $x W(x) \kappa(x, \varphi)$ vanishes for $x \rightarrow 0$ and $x \rightarrow \infty$. This can be achieved by picking a weighting function which drops off sufficiently fast. Expressing the radial derivative in Cartesian coordinates and using $\mathcal{F} = \nabla \kappa$ (Bacon et al. 2006), we can write the mass moments in terms of first flexion:

$$M_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi \mathcal{F}_t(x, \varphi), \quad (\text{A5})$$

where

$$\mathcal{F}_t(x, \varphi) = -[\mathcal{F}_1(x, \varphi) \cos(\varphi) + \mathcal{F}_2(x, \varphi) \sin(\varphi)], \quad (\text{A6})$$

$$F_t(x, \varphi) = -[F_1(x, \varphi) \cos(\varphi) + F_2(x, \varphi) \sin(\varphi)], \quad (\text{A7})$$

defines the non-reduced and reduced tangential first flexion. We use a mass-sheet reconstruction κ_{rec} to find

$$M_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}(x, \varphi)) F_t(x, \varphi). \quad (\text{A8})$$

Note that κ_{rec} is different from the convergence that is used to calculate the moments in equations (A1) and (A2). Applying the same operations to equation (A2), we can also express the multipole moments using first flexion:

$$Q_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} \mathcal{F}_t(x, \varphi) \quad (\text{A9})$$

$$= \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi)) F_t(x, \varphi). \quad (\text{A10})$$

The moments can also be expressed using a combination of \mathcal{F} and \mathcal{G} . We define the tangential and radial shear and the corresponding reduced quantities by

$$\gamma_t(x, \varphi) = -[\gamma_1(x, \varphi) \cos(2\varphi) + \gamma_2(x, \varphi) \sin(2\varphi)], \quad (\text{A11})$$

$$\gamma_r(x, \varphi) = -[\gamma_2(x, \varphi) \cos(2\varphi) - \gamma_1(x, \varphi) \sin(2\varphi)], \quad (\text{A12})$$

$$g_t(x, \varphi) = -[g_1(x, \varphi) \cos(2\varphi) + g_2(x, \varphi) \sin(2\varphi)], \quad (\text{A13})$$

$$g_r(x, \varphi) = -[g_2(x, \varphi) \cos(2\varphi) - g_1(x, \varphi) \sin(2\varphi)], \quad (\text{A14})$$

and use the relation shown in Kaiser (1995)

$$\nabla \kappa = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}. \quad (\text{A15})$$

Inserting the radial part of this relation transformed to polar coordinates into equation (A3), using trigonometric identities,

integrating by parts and demanding that the surface terms $xW(x)\gamma_t(x, \varphi)$ vanish for $x \rightarrow 0, \infty$ leads to an expression of the mass moments using shear (SB97),

$$M_\gamma^{(n)} = \int_0^\infty dx [2W(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi \gamma_t(x, \varphi) = \int_0^\infty dx [2W(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}(x, \varphi))g_t(x, \varphi). \quad (\text{A16})$$

We define

$$V(x) = \int_0^x dy W(y) \quad (\text{A17})$$

and require that $[2V(x) - xW(x)]\gamma_t(x, \varphi) \rightarrow 0$ for $x \rightarrow 0, x \rightarrow \infty$. Then the boundary terms resulting from a partial integration with respect to x vanish and we find

$$M_\gamma^{(n)} = - \int_0^\infty dx [2V(x) - xW(x)] \int_0^{2\pi} d\varphi \frac{\partial \gamma_t(x, \varphi)}{\partial x}. \quad (\text{A18})$$

We further define the radial first flexion, the tangential and radial second flexion and their reduced counterparts by

$$\mathcal{F}_r(x, \varphi) = -[\mathcal{F}_2(x, \varphi) \cos(\varphi) - \mathcal{F}_1(x, \varphi) \sin(\varphi)], \quad (\text{A19})$$

$$\mathcal{G}_t(x, \varphi) = -[\mathcal{G}_1(x, \varphi) \cos(3\varphi) + \mathcal{G}_2(x, \varphi) \sin(3\varphi)], \quad (\text{A20})$$

$$\mathcal{G}_r(x, \varphi) = -[\mathcal{G}_2(x, \varphi) \cos(3\varphi) - \mathcal{G}_1(x, \varphi) \sin(3\varphi)], \quad (\text{A21})$$

$$F_r(x, \varphi) = -[F_2(x, \varphi) \cos(\varphi) - F_1(x, \varphi) \sin(\varphi)], \quad (\text{A22})$$

$$G_t(x, \varphi) = -[G_1(x, \varphi) \cos(3\varphi) + G_2(x, \varphi) \sin(3\varphi)], \quad (\text{A23})$$

$$G_r(x, \varphi) = -[G_2(x, \varphi) \cos(3\varphi) - G_1(x, \varphi) \sin(3\varphi)]. \quad (\text{A24})$$

Now we can express the radial derivative in Cartesian coordinates, apply trigonometric identities and use the relations between flexion and shear derivatives in equation (11) to obtain the mass moments in terms of first and second flexion:

$$M_{\mathcal{F}, \mathcal{G}}^{(n)} = - \int_0^\infty dx [V(x) - \frac{1}{2}xW(x)] \int_0^{2\pi} d\varphi [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] \quad (\text{A25})$$

$$= - \int_0^\infty dx [V(x) - \frac{1}{2}xW(x)] \int_0^{2\pi} d\varphi [(1 - \kappa_{\text{rec}}(x, \varphi))F_t(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_t(x, \varphi)]. \quad (\text{A26})$$

The multipole moments using \mathcal{F} and \mathcal{G} can be derived in an analogous way. We integrate equation (A2) by parts and require that the boundary term $xW(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow 0$ and $x \rightarrow \infty$. Inserting the radial part of equation (A15) transformed to polar coordinates, integrating by parts and demanding that $xW(x)\gamma_t(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$ leads to the multipole moments in terms of shear (SB97),

$$Q_\gamma^{(n)} = \int_0^\infty dx [2W(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi) - in \int_0^\infty dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_r(x, \varphi) \quad (\text{A27})$$

$$= \int_0^\infty dx [2W(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi))g_t(x, \varphi) - in \int_0^\infty dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi))g_r(x, \varphi). \quad (\text{A28})$$

Partial integration with respect to x and applying the same transformations and identities as before leads to

$$Q_{\mathcal{F}, \mathcal{G}}^{(n)} = - \int_0^\infty dx [V(x) - \frac{1}{2}xW(x)] \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] \\ + i \frac{n}{2} \int_0^\infty dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_r(x, \varphi) + \mathcal{G}_r(x, \varphi)] \quad (\text{A29})$$

$$= - \int_0^\infty dx [V(x) - \frac{1}{2}xW(x)] \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}(x, \varphi))F_t(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_t(x, \varphi)] \\ + i \frac{n}{2} \int_0^\infty dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}(x, \varphi))F_r(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_r(x, \varphi)], \quad (\text{A30})$$

where we required that

$$[2V(x) - xW(x)]\gamma_t(x, \varphi) \rightarrow 0 \text{ for } x \rightarrow 0, x \rightarrow \infty, \quad (\text{A31})$$

$$V(x)\gamma_r(x, \varphi) \rightarrow 0 \text{ for } x \rightarrow 0, x \rightarrow \infty, \quad (\text{A32})$$

so that the boundary terms in the partial integration vanish.

A1.1 Moments on rings

We now extend our moment formulas to rings, i.e. we omit the innermost part of the lens in the integration, because we typically do not have weak lensing information in this area. The derivation is exactly analogous to the one presented in the previous section and we will therefore only state the resulting equivalent formulas and the respective requirements for the weighting function. Furthermore we will not explicitly denote the reconstruction of shear and flexion using κ_{rec} to keep this section brief. We define

$$M_{R;\kappa}^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \kappa(x, \varphi) \quad (\text{A33})$$

and the corresponding multipole moments

$$Q_{R;\kappa}^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa(x, \varphi). \quad (\text{A34})$$

We further define

$$xW_R(x) = \int_R^x dy y^{n+1} w(y) \quad (\text{A35})$$

and demand that $xW_R(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. Note that this is simply a different choice of the integration constant. We obtain

$$M_{R;\mathcal{F}}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi \mathcal{F}_t(x, \varphi), \quad (\text{A36})$$

$$Q_{R;\mathcal{F}}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \mathcal{F}_t(x, \varphi). \quad (\text{A37})$$

We further demand that the term $xW_R(x)\gamma_t(x, \varphi)$ vanishes for $x \rightarrow R, \infty$ and note that the other surface terms cancel due to periodicity. We have now

$$M_{R;\gamma}^{(n)} = \int_R^\infty dx [2W_R(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi \gamma_t(x, \varphi), \quad (\text{A38})$$

$$Q_{R;\gamma}^{(n)} = \int_R^\infty dx [2W_R(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi) - i n \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_r(x, \varphi). \quad (\text{A39})$$

We define

$$V_R(x) = \int_R^x dy W_R(y) \quad (\text{A40})$$

and require that $[2V_R(x) - xW_R(x)]\gamma_t(x, \varphi)$ and, for the multipole moments, $V_R(x)\gamma_r(x, \varphi)$ vanish for $x \rightarrow R, \infty$ to find

$$M_{R;\mathcal{F},\mathcal{G}}^{(n)} = - \int_R^\infty dx [V_R(x) - \frac{1}{2}x W_R(x)] \int_0^{2\pi} d\varphi [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)], \quad (\text{A41})$$

$$\begin{aligned} Q_{R;\mathcal{F},\mathcal{G}}^{(n)} = & - \int_R^\infty dx [V_R(x) - \frac{1}{2}x W_R(x)] \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] \\ & + i \frac{n}{2} \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_r(x, \varphi) + \mathcal{G}_r(x, \varphi)]. \end{aligned} \quad (\text{A42})$$

If weak lensing data is only available in a small part of the ring, we can further restrict the integration to a partial ring. For this purpose, we modify the mass and multipole moment equations for κ and \mathcal{F} by replacing 2π with the desired maximum angle ϕ and their equivalence still holds. However, they are now no longer equivalent to the γ and \mathcal{F}, \mathcal{G} moments. The integration area of these moments cannot be restricted in general, since the derivation of the equivalences uses the periodicity due to the 2π boundary. However, in the case of the mass moments $M_{R;\gamma}^{(n)}$ and $M_{R;\mathcal{F},\mathcal{G}}^{(n)}$, replacing 2π with π also preserves the required periodicity and thus it is possible to restrict these moments to one half of a ring. Therefore all mass moments are also equivalent for $\phi = \pi$.

APPENDIX B: DERIVATION OF THE MASS AND MULTIPOLE MOMENT EQUIVALENCES FOR K , REDUCED SHEAR, AND REDUCED FLEXIONS

We derive the moment equivalence relations for the reduced quantities. Following [Cain et al. \(2011\)](#), we define

$$K(x, \varphi, x_0, \varphi_0) = -\ln(1 - \kappa(x, \varphi, x_0, \varphi_0)) \quad (\text{B1})$$

and thus

$$K_{,i} = \frac{1}{1 - \kappa} \kappa_{,i} = F_i. \quad (\text{B2})$$

We want $\kappa < 1$ and choose the lower integral limit for our ring, R , such that $\kappa(x, \varphi) < 1 \forall x \geq R, \forall \varphi$. Therefore $K(x, \varphi)$ is well-defined and finite and we can define

$$M_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi K(x, \varphi, x_0, \varphi_0), \quad (\text{B3})$$

$$Q_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} K(x, \varphi, x_0, \varphi_0). \quad (\text{B4})$$

In the following, we will again not explicitly denote x_0 and φ_0 to keep the notation simple. We require that $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$ and integrate equation (B3) by parts with respect to x to find

$$M_K^{(n)} = - \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi \frac{\partial K(x, \varphi)}{\partial x}. \quad (\text{B5})$$

We now express the radial derivative in Cartesian coordinates, apply equation (B2) and use the definition of F_t in equation (A7). Then we have

$$M_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi F_t(x, \varphi). \quad (\text{B6})$$

In the exact same way we find for the multipole moments

$$Q_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} F_t(x, \varphi). \quad (\text{B7})$$

Now we would like to express the moments using the reduced shear. Therefore we rewrite the K derivatives by using equation (A15) and the definition of g :

$$K_{,1} = g_{1,1} - g_1 K_{,1} + g_{2,2} - g_2 K_{,2}, \quad (\text{B8})$$

$$K_{,2} = g_{2,1} - g_2 K_{,1} - g_{1,2} + g_1 K_{,2}. \quad (\text{B9})$$

Furthermore, we use $K_{,i} = F_i$ to find

$$\frac{\partial K}{\partial x} = -F_t, \quad (\text{B10})$$

$$\frac{\partial K}{\partial \varphi} = -x F_r. \quad (\text{B11})$$

Transforming the Cartesian derivatives on the right hand side of equations (B8) and (B9) into polar coordinates, using trigonometric identities and the definitions of the tangential and radial reduced shear and reduced flexion, we have

$$\frac{\partial K}{\partial x} = \cos(\varphi) K_{,1} + \sin(\varphi) K_{,2} = -\left[\frac{\partial g_t}{\partial x} - \frac{\cos(2\varphi)}{x} \frac{\partial g_2}{\partial \varphi} + \frac{\sin(2\varphi)}{x} \frac{\partial g_1}{\partial \varphi} + g_t F_t + g_r F_r\right]. \quad (\text{B12})$$

We insert this expression into equation (B5), integrate the term with the x -derivative by parts and demand that $xW_R(x)g_t(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$. Then the surface term vanishes and we have

$$\begin{aligned} M_F^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t F_t + g_r F_r) - \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi g_t \\ & - \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi (\cos(2\varphi) \frac{\partial g_2}{\partial \varphi} - \sin(2\varphi) \frac{\partial g_1}{\partial \varphi}). \end{aligned} \quad (\text{B13})$$

Integrating the φ derivatives by parts and noting that the surface terms vanish due to their periodicity, we find

$$\begin{aligned} M_g^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) \\ & + \int_R^\infty dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi g_t(x, \varphi). \end{aligned} \quad (\text{B14})$$

Using the same steps, we can derive the multipole moments in terms of reduced shear and reduced flexion,

$$Q_g^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) \\ + \int_R^\infty dx (2W_R(x) - x^{n+1}w(x)) \int_0^{2\pi} d\varphi e^{in\varphi} g_t(x, \varphi) - in \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} g_r(x, \varphi). \quad (\text{B15})$$

We can integrate the second term in equation (B14) by parts with respect to x and demand that the surface term $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$. We have now

$$M_g^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) \\ - \int_R^\infty dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi \frac{\partial g_t(x, \varphi)}{\partial x}. \quad (\text{B16})$$

Again applying a transformation to Cartesian coordinates, trigonometric identities, $\nabla\kappa = \mathcal{F}$, and the relations from equation (11), we find for the derivative

$$\frac{\partial g_t}{\partial x} = \frac{1}{2}(F_t + G_t) - g_t F_t \quad (\text{B17})$$

and thus we have for the mass moment

$$M_{F,G}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_r(x, \varphi) F_r(x, \varphi) + \frac{1}{2}(F_t(x, \varphi) + G_t(x, \varphi))) \\ - \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi (F_t(x, \varphi) + G_t(x, \varphi) - 2g_t(x, \varphi) F_t(x, \varphi)). \quad (\text{B18})$$

Analogously, we can integrate the last two terms in equation (B15) by parts and demand that the resulting surface terms $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ and $V_R(x)g_r(x, \varphi)$ vanish for $x \rightarrow R$ and $x \rightarrow \infty$. In the same way as above, we can calculate the derivative as

$$\frac{\partial g_r}{\partial x} = \frac{1}{2}(F_r + G_r) - g_r F_t \quad (\text{B19})$$

and use our result for $\partial g_t/\partial x$ to find

$$Q_{F,G}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) - \int_R^\infty dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi e^{in\varphi} \\ \cdot \left(\frac{1}{2}(F_t(x, \varphi) + G_t(x, \varphi)) - g_t(x, \varphi) F_t(x, \varphi) \right) + in \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \\ \cdot \left(\frac{1}{2}(F_r(x, \varphi) + G_r(x, \varphi)) - g_r(x, \varphi) F_t(x, \varphi) \right). \quad (\text{B20})$$

If weak lensing data is only available in a small angular window, we can restrict the integration area to a partial ring. We modify the mass and multipole moment equations for K and F by replacing 2π with the maximum angle ϕ and they will still be equivalent. However, they are then no longer equivalent to the g and F, G moments. The integration area of these moments cannot be further restricted, because otherwise the cancellation of the surface terms due to the periodicity would no longer hold.

APPENDIX C: MASS MOMENT DISPERSIONS

C1 Shear, flexion, and convergence moments

We calculate the mass moment dispersions in the absence of lensing. For the convergence, we have $\langle M_\kappa^{(n)} \rangle = 0$ and

$$\sigma_{M,\kappa}^2 = \langle |M_\kappa^{(n)}|^2 \rangle \quad (\text{C1})$$

$$= \int_0^\infty \int_0^\infty dx dy x^{n+1} y^{n+1} w(x) w(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \kappa(x, \varphi) \kappa(y, \varphi') \rangle. \quad (\text{C2})$$

Following Schneider et al. (2000), the magnification signal for number counts is $|\Delta n_\kappa| = |\mu^{\beta-1} - 1| n_\kappa$, where μ is the magnification and n_κ is the source density. The Poisson noise is $\sqrt{N_{\kappa,a}}$, where $N_{\kappa,a}$ is the number of sources over which we average to determine κ in a given area a . We neglect the additional noise due to source clustering. In the weak lensing regime and assuming $\beta = 0.5$, we have $\mu \approx 1 + 2\kappa$ and thus $|\kappa| \approx |\Delta n_\kappa|/n_\kappa$.

We assume a discrete distribution of κ over an area A and measure the convergence by averaging our source counts

in an area a . Therefore our measurement points of κ are uncorrelated. Thus we have for the sum of the convergence variance over the area A :

$$\frac{n_\kappa^2 a^2}{n_\kappa^2} \sum_{i=1}^{N_\kappa/n_\kappa a} \sum_{j=1}^{N_\kappa/n_\kappa a} \langle \kappa_i \kappa_j \rangle = a^2 \sum_{i=1}^{N_\kappa/n_\kappa a} \frac{1}{n_\kappa a} = \frac{A}{n_\kappa}, \quad (\text{C3})$$

$$\langle \kappa_i \kappa_j \rangle = \frac{1}{n_\kappa a} \delta_{ij}. \quad (\text{C4})$$

Extending this to integration, we have

$$\int \int dx^2 dy^2 \langle \kappa(\vec{x}) \kappa(\vec{y}) \rangle = \frac{1}{n_\kappa} \int dx^2 = \frac{A}{n_\kappa} \quad (\text{C5})$$

and thus

$$\langle \kappa(\vec{x}) \kappa(\vec{y}) \rangle = \frac{1}{n_\kappa} \delta(\vec{x} - \vec{y}) \quad (\text{C6})$$

with the Dirac delta distribution $\delta(\vec{x} - \vec{y})$.

Therefore we have

$$\sigma_{M,\kappa}^2 = \frac{2\pi}{n_\kappa} \int_0^\infty dx x^{2n+1} w(x)^2 \quad (\text{C7})$$

and thus the standard deviation

$$\sigma_{M,\kappa} = \sqrt{\frac{2\pi}{n_\kappa}} \left(\int_0^\infty dx x^{2n+1} w(x)^2 \right)^{\frac{1}{2}}. \quad (\text{C8})$$

Before calculating the shear and flexion moments, we make the following observation. In the case of a discrete distribution over an area A and no lensing, all our measurements of the shear are uncorrelated and we have

$$\frac{1}{n_\gamma^2} \sum_{i=1}^N \sum_{j=1}^N \langle \gamma_i \gamma_j \rangle = \frac{1}{n_\gamma^2} \sum_{i=1}^N \sigma_\gamma^2 = \frac{A}{n_\gamma} \sigma_\gamma^2, \quad (\text{C9})$$

$$\langle \gamma_i \gamma_j \rangle = \sigma_\gamma^2 \delta_{ij} \quad (\text{C10})$$

with the number density $n_\gamma = N_\gamma/A$, where N_γ is the number of sources where we measured γ . Extending this to integration, we have

$$\int \int dx^2 dy^2 \langle \gamma(\vec{x}) \gamma(\vec{y}) \rangle = \frac{1}{n_\gamma} \int dx^2 \sigma_\gamma^2 = \frac{A}{n_\gamma} \sigma_\gamma^2 \quad (\text{C11})$$

and thus

$$\langle \gamma(\vec{x}) \gamma(\vec{y}) \rangle = \frac{\sigma_\gamma^2}{n_\gamma} \delta(\vec{x} - \vec{y}) \quad (\text{C12})$$

with the Dirac delta distribution $\delta(\vec{x} - \vec{y})$.

The errors on κ_{rec} and g_t generally depend on the measurement technique. However, we are assuming the absence of a lens for the calculation of the standard deviation, so we can look at the mass moment in terms of γ instead of κ_{rec} and g_t and we neglect the measurement noise. Therefore we have $\langle M_\gamma^{(n)} \rangle = 0$ and for the variance:

$$\sigma_{M,\gamma}^2 = \langle |M_\gamma^{(n)}|^2 \rangle \quad (\text{C13})$$

$$= \int_0^\infty \int_0^\infty dx dy [2W(x) - x^{n+1}w(x)][2W(y) - y^{n+1}w(y)] \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \gamma_t(x, \varphi) \gamma_t(y, \varphi') \rangle. \quad (\text{C14})$$

We use

$$\langle \gamma_t(x, \varphi) \gamma_t(y, \varphi') \rangle = \frac{\sigma_\gamma^2}{n_\gamma} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C15})$$

where we have written the Dirac delta distribution in polar coordinates, and get

$$\sigma_{M,\gamma}^2 = \int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \int_0^{2\pi} d\varphi \frac{\sigma_\gamma^2}{n_\gamma} \quad (\text{C16})$$

and for the standard deviation

$$\sigma_{M,\gamma} = \sqrt{\frac{2\pi}{n_\gamma}} \sigma_\gamma \left(\int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \right)^{\frac{1}{2}}. \quad (\text{C17})$$

Similarly, we have $\langle M_{\mathcal{F}}^{(n)} \rangle = 0$ and

$$\sigma_{M,\mathcal{F}}^2 = \langle |M_{\mathcal{F}}^{(n)}|^2 \rangle \quad (\text{C18})$$

$$= \int_0^\infty \int_0^\infty dx dy xW(x)yW(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \mathcal{F}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle. \quad (\text{C19})$$

Using

$$\langle \mathcal{F}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi'), \quad (\text{C20})$$

we have

$$\sigma_{M,\mathcal{F}}^2 = \int_0^\infty dx xW(x)^2 \int_0^{2\pi} d\varphi \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \quad (\text{C21})$$

and consequently

$$\sigma_{M,\mathcal{F}} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}}} \sigma_{\mathcal{F}} \left(\int_0^\infty dx xW(x)^2 \right)^{\frac{1}{2}}. \quad (\text{C22})$$

In the same way, we get $\langle M_{\mathcal{F},\mathcal{G}}^{(n)} \rangle = 0$ and

$$\sigma_{M,\mathcal{F},\mathcal{G}}^2 = \langle |M_{\mathcal{F},\mathcal{G}}^{(n)}|^2 \rangle \quad (\text{C23})$$

$$= \int_0^\infty \int_0^\infty dx dy [V(x) - \frac{1}{2}xW(x)][V(y) - \frac{1}{2}yW(y)] \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \mathcal{F}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle \quad (\text{C24})$$

$$+ \langle \mathcal{F}_t(x, \varphi) \mathcal{G}_t(y, \varphi') \rangle + \langle \mathcal{G}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle + \langle \mathcal{G}_t(x, \varphi) \mathcal{G}_t(y, \varphi') \rangle.$$

We use equation (C20) and

$$\langle \mathcal{G}_t(x, \varphi) \mathcal{G}_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi') \quad (\text{C25})$$

and assume that

$$\langle \mathcal{F}_t(x, \varphi) \mathcal{G}_t(y, \varphi') \rangle = 0 \quad (\text{C26})$$

to get

$$\sigma_{M,\mathcal{F},\mathcal{G}}^2 = \int_0^\infty dx \frac{1}{x} [V(x) - \frac{1}{2}xW(x)]^2 \int_0^{2\pi} d\varphi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \right) \quad (\text{C27})$$

and the standard deviation

$$\sigma_{M,\mathcal{F},\mathcal{G}} = \sqrt{2\pi} \sqrt{\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}}} \left(\int_0^\infty dx \frac{1}{x} [V(x) - \frac{1}{2}xW(x)]^2 \right)^{\frac{1}{2}}. \quad (\text{C28})$$

C2 Reduced shear, reduced flexion, and K moments

We have

$$\langle M_K^{(n)} \rangle = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \langle K(x, \varphi) \rangle. \quad (\text{C29})$$

In the absence of lensing, κ is typically everywhere smaller than 1, so we can use the Mercator series,

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \text{ for } |x| < 1, \quad (\text{C30})$$

to find

$$\langle K(x, \varphi) \rangle = (-1) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (-1)^n \langle \kappa(x, \varphi)^n \rangle. \quad (\text{C31})$$

As κ is small, we can restrict ourselves to the 2 lowest order terms,

$$\langle \kappa(x, \varphi) \rangle = 0, \quad (\text{C32})$$

$$\langle \kappa(x, \varphi) \kappa(x, \varphi) \rangle = \frac{1}{n_\kappa a}, \quad (\text{C33})$$

where we again average the sources over the area a , and get

$$\langle K(x, \varphi) \rangle = \frac{1}{2n_\kappa a}. \quad (\text{C34})$$

Therefore we have

$$\langle M_K^{(n)} \rangle = \frac{\pi}{n_\kappa a} \int_R^\infty dx x^{n+1} w(x). \quad (\text{C35})$$

The variance is

$$\sigma_{M,K}^2 = \langle |M_K^{(n)}|^2 \rangle - \langle M_K^{(n)} \rangle^2. \quad (\text{C36})$$

We have

$$\langle |M_K^{(n)}|^2 \rangle = \int_R^\infty \int_R^\infty dx dy x^{n+1} y^{n+1} w(x) w(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle K(x, \varphi) K(y, \varphi') \rangle \quad (\text{C37})$$

and

$$\langle K(x, \varphi) K(y, \varphi') \rangle = \langle \ln(1 - \kappa(x, \varphi)) \ln(1 - \kappa(y, \varphi')) \rangle \quad (\text{C38})$$

and using again the Mercator series, we get

$$\langle K(x, \varphi) K(y, \varphi') \rangle = \left\langle \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (-1)^n \kappa^n(x, \varphi) \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} (-1)^m \kappa^m(y, \varphi') \right\rangle \quad (\text{C39})$$

$$= \left\langle \sum_{n=1}^{\infty} \frac{\kappa^n(x, \varphi)}{n} \sum_{m=1}^{\infty} \frac{\kappa^m(y, \varphi')}{m} \right\rangle. \quad (\text{C40})$$

As κ is small, we can ignore the higher order terms to get

$$\langle K(x, \varphi) K(y, \varphi') \rangle \approx \langle \kappa(x, \varphi) \kappa(y, \varphi') \rangle = \frac{1}{n_\kappa} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi'). \quad (\text{C41})$$

Thus we have

$$\sigma_{M,K}^2 = \frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 a^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2, \quad (\text{C42})$$

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 a^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}. \quad (\text{C43})$$

The variance is typically well behaved. However, as the first term depends on the source density and the second on the number of sources over which we average to obtain the convergence, $n_\kappa a$, it is theoretically possible to construct unreasonable combinations. For example, obtaining the convergence from only 1 source per bin while having a high source density would lead to an unreasonable result, as the contribution from the poorly constrained expectation value would dominate the other uncertainties. Naturally, such a combination would be avoided in real applications.

As we treat the case of no lensing, $\langle \kappa \rangle = 0$ and we can thus average over the whole area A to obtain our convergence estimate. Thus we have

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 A^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}. \quad (\text{C44})$$

We have $\langle M_g^{(n)} \rangle = 0$ and

$$\sigma_{M,g}^2 = \langle |M_g^{(n)}|^2 \rangle \quad (\text{C45})$$

$$= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \quad (\text{C46})$$

$$\begin{aligned} & (\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle + \langle g_t(x, \varphi) F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle \\ & + \langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle + \langle g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle) \\ & + \int_R^\infty \int_R^\infty dx dy x W_R(x) (2W_R(y) - y^{n+1} w(y)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \\ & (\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') \rangle + \langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') \rangle) \\ & + \int_R^\infty \int_R^\infty dx dy y W_R(y) (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \\ & (\langle g_t(y, \varphi') F_t(y, \varphi') g_t(x, \varphi) \rangle + \langle g_r(y, \varphi') F_r(y, \varphi') g_t(x, \varphi) \rangle) \\ & + \int_R^\infty \int_R^\infty dx dy (2W_R(x) - x^{n+1} w(x)) (2W_R(y) - y^{n+1} w(y)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle g_t(x, \varphi) g_t(y, \varphi') \rangle. \end{aligned}$$

In the absence of lensing, we expect the flexion and the shear to be uncorrelated. Therefore we have, using the relation shown in Goodman (1960),

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle \quad (\text{C47})$$

$$= (\langle g_t(x, \varphi) g_t(x, \varphi) \rangle \langle F_t(x, \varphi) F_t(x, \varphi) \rangle + \langle g_t(x, \varphi) g_t(x, \varphi) \rangle \langle F_t(x, \varphi) \rangle^2) \quad (\text{C48})$$

$$+ \langle F_t(x, \varphi) F_t(x, \varphi) \rangle \langle g_t(x, \varphi) \rangle^2) \frac{1}{n_{\gamma, \mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi') \quad (\text{C49})$$

$$= \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\gamma, \mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C50})$$

where $n_{\gamma, \mathcal{F}}$ is the number density of sources for which we have both shear and flexion information. As flexion is typically much harder to measure than shear, we can make the approximation $n_{\gamma, \mathcal{F}} \approx n_{\mathcal{F}}$. Using this approximation and making a similar calculation in the other cases, we have

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle = \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C51})$$

$$\langle g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle = \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C52})$$

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle = \langle g_t(x, \varphi) \rangle \langle g_r(y, \varphi') \rangle \langle F_t(x, \varphi) \rangle \langle F_r(y, \varphi') \rangle = 0, \quad (\text{C53})$$

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') \rangle = \langle g_t(x, \varphi) g_t(y, \varphi') \rangle \langle F_t(x, \varphi) \rangle = 0, \quad (\text{C54})$$

$$\langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') \rangle = \langle g_r(x, \varphi) g_t(y, \varphi') \rangle \langle F_r(x, \varphi) \rangle = 0, \quad (\text{C55})$$

and thus

$$\sigma_{M,g}^2 = \int_R^\infty dx x W_R(x)^2 \int_0^{2\pi} d\varphi 2 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1} w(x))^2 \int_0^{2\pi} d\varphi \frac{\sigma_\gamma^2}{n_\gamma} \quad (\text{C56})$$

and we have the standard deviation

$$\sigma_{M,g} = \left(\frac{4\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \sigma_\gamma^2 \left(\int_R^\infty dx x W_R(x)^2 \right) + \frac{2\pi}{n_\gamma} \sigma_\gamma^2 \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1} w(x))^2 \right)^{\frac{1}{2}}. \quad (\text{C57})$$

We have $\langle M_F^{(n)} \rangle = 0$ and

$$\sigma_{M,F}^2 = \langle |M_F^{(n)}|^2 \rangle \quad (\text{C58})$$

$$= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle F_t(x, \varphi) F_t(y, \varphi') \rangle \quad (\text{C59})$$

and using

$$< F_t(x, \varphi) F_t(y, \varphi') > = \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi') \quad (\text{C60})$$

we have

$$\sigma_{M,F}^2 = \frac{2\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \int_R^\infty dx x W_R(x)^2 \quad (\text{C61})$$

and the standard deviation

$$\sigma_{M,F} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \left(\int_R^\infty dx x W_R(x)^2 \right)^{\frac{1}{2}}}. \quad (\text{C62})$$

We have $< M_{F,G}^{(n)} > = 0$ and

$$\sigma_{M,F,G}^2 = < |M_{F,G}^{(n)}|^2 > \quad (\text{C63})$$

$$\begin{aligned} &= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (< g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') > \\ &+ \frac{1}{2} < g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') > + \frac{1}{2} < g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') > + \frac{1}{2} < F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') > \\ &+ \frac{1}{2} < G_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') > + \frac{1}{4} < F_t(x, \varphi) F_t(y, \varphi') > + \frac{1}{4} < F_t(x, \varphi) G_t(y, \varphi') > \\ &+ \frac{1}{4} < G_t(x, \varphi) F_t(y, \varphi') > + \frac{1}{4} < G_t(x, \varphi) G_t(y, \varphi') >) \\ &- \int_R^\infty \int_R^\infty dx dy x W_R(x) V_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (< g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') > \\ &+ < g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') > - 2 < g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > \\ &+ \frac{1}{2} < F_t(x, \varphi) F_t(y, \varphi') > + \frac{1}{2} < F_t(x, \varphi) G_t(y, \varphi') > - < F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > \\ &+ \frac{1}{2} < G_t(x, \varphi) F_t(y, \varphi') > + \frac{1}{2} < G_t(x, \varphi) G_t(y, \varphi') > - < G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') >) \\ &- \int_R^\infty \int_R^\infty dx dy V_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (< g_r(y, \varphi') F_r(y, \varphi') F_t(x, \varphi) > \\ &+ < g_r(y, \varphi') F_r(y, \varphi') G_t(x, \varphi) > - 2 < g_r(y, \varphi') F_r(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) > \\ &+ \frac{1}{2} < F_t(y, \varphi') F_t(x, \varphi) > + \frac{1}{2} < F_t(y, \varphi') G_t(x, \varphi) > - < F_t(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) > \\ &+ \frac{1}{2} < G_t(y, \varphi') F_t(x, \varphi) > + \frac{1}{2} < G_t(y, \varphi') G_t(x, \varphi) > - < G_t(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) >) \\ &+ \int_R^\infty \int_R^\infty dx dy V_R(x) V_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (< F_t(x, \varphi) F_t(y, \varphi') > + < F_t(x, \varphi) G_t(y, \varphi') > \\ &- 2 < F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > + < G_t(x, \varphi) F_t(y, \varphi') > + < G_t(x, \varphi) G_t(y, \varphi') > \\ &- 2 < G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > - 2 < g_t(x, \varphi) F_t(x, \varphi) F_t(y, \varphi') > \\ &- 2 < g_t(x, \varphi) F_t(x, \varphi) G_t(y, \varphi') > + 4 < g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') >). \end{aligned} \quad (\text{C64})$$

We again assume that the first and second flexions and the shear are uncorrelated in the absence of lensing, thus we can use the previously derived relations and we also have

$$< G_t(x, \varphi) G_t(y, \varphi') > = \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C65})$$

$$< g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') > = < g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') > = < F_t(x, \varphi) G_t(y, \varphi') > = 0, \quad (\text{C66})$$

$$< F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > = < g_t(y, \varphi') > < F_t(x, \varphi) F_t(y, \varphi') > = 0, \quad (\text{C67})$$

$$< G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') > = 0. \quad (\text{C68})$$

This gives us

$$\begin{aligned} \sigma_{M,F,G}^2 = 2\pi \left(\frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{F}}^2}{4n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{4n_{\mathcal{G}}} \right) \int_R^\infty dx \, x W_R(x)^2 - 2\pi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \right) \int_R^\infty dx \, W_R(x) V_R(x) \\ + 2\pi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} + 4 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \right) \int_R^\infty dx \, \frac{1}{x} V_R(x)^2, \end{aligned} \quad (\text{C69})$$

$$\begin{aligned} \sigma_{M,F,G} = \sqrt{2\pi} \left(\left(\frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{F}}^2}{4n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{4n_{\mathcal{G}}} \right) \int_R^\infty dx \, x W_R(x)^2 - \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \right) \int_R^\infty dx \, W_R(x) V_R(x) \right. \\ \left. + \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} + 4 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \right) \int_R^\infty dx \, \frac{1}{x} V_R(x)^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (\text{C70})$$

APPENDIX D: CONVERGENCE, SHEAR, AND FLEXION FOR NFW PROFILES

Navarro, Frenk, and White showed that spherically averaged cold dark matter (CDM) halo density profiles can be fitted over two decades in radius by the NFW density profile (Navarro et al. 1996, 1997)

$$\frac{\rho(r)}{\rho_{\text{crit}}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2}, \quad (\text{D1})$$

where r_s is a scale radius, δ_c is a characteristic dimensionless density, and $\rho_{\text{crit}} = 3H^2/8\pi G$ is the critical density for closure. This profile leads to the convergence formula (Bartelmann 1996; Wright & Brainerd 2000)

$$\kappa_{\text{NFW}}(y) = \frac{2\rho_{\text{crit}} \delta_c r_s}{\Sigma_{\text{crit}}} \frac{f(y)}{y^2 - 1}, \quad (\text{D2})$$

where $y = \xi/r_s$ and

$$f(y) = \begin{cases} 1 - \frac{2}{\sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right), & y < 1 \\ \frac{y^2-1}{3}, & y = 1 \\ 1 - \frac{2}{\sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right), & y > 1. \end{cases} \quad (\text{D3})$$

The NFW profile is spherically symmetric and therefore $\gamma_{r,\text{NFW}} = 0$ (see e.g. Meylan et al. 2006). The tangential shear is (Wright & Brainerd 2000)

$$\gamma_{t,\text{NFW}} = \frac{\rho_{\text{crit}} \delta_c r_s}{\Sigma_{\text{crit}}} g(y), \quad (\text{D4})$$

where

$$g(y) = \begin{cases} \frac{8}{y^2 \sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right) + \frac{4}{y^2} \log \left(\frac{y}{2} \right) - \frac{2}{y^2-1} + \frac{4}{(y^2-1) \sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right), & y < 1 \\ \frac{10}{3} + 4 \log \left(\frac{1}{2} \right), & y = 1 \\ \frac{8}{y^2 \sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right) + \frac{4}{y^2} \log \left(\frac{y}{2} \right) - \frac{2}{y^2-1} + \frac{4}{(y^2-1)^{3/2}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right), & y > 1. \end{cases} \quad (\text{D5})$$

We can use the expressions for first and second flexion for a NFW profile derived in Bacon et al. (2006) to obtain

$$\mathcal{F}_{t,\text{NFW}} = \frac{2\rho_{\text{crit}} \delta_c D_{\text{ol}}}{\Sigma_{\text{crit}} (y^2 - 1)^2} \left[2y f(y) - h(y) \right], \quad (\text{D6})$$

$$\mathcal{G}_{t,\text{NFW}} = -\frac{2\rho_{\text{crit}} \delta_c D_{\text{ol}}}{\Sigma_{\text{crit}}} \left[\frac{8}{y^3} \log \left(\frac{y}{2} \right) + \frac{(3(1-2y^2)/y) + l(y)}{(y^2 - 1)^2} \right], \quad (\text{D7})$$

where

$$h(y) = \begin{cases} \frac{2y}{\sqrt{1-y^2}} \operatorname{arctanh}\left(\sqrt{\frac{1-y}{1+y}}\right) - \frac{1}{y}, & y < 1 \\ \frac{2y(y^2-1)}{3} - 0.4(y^2-1)^2, & y = 1 \\ \frac{2y}{\sqrt{y^2-1}} \operatorname{arctan}\left(\sqrt{\frac{y-1}{y+1}}\right) - \frac{1}{y}, & y > 1, \end{cases} \quad (\text{D8})$$

$$l(y) = \begin{cases} \left(\frac{8}{y^3} - \frac{20}{y} + 15y\right) \frac{2}{\sqrt{1-y^2}} \operatorname{arctanh}\left(\sqrt{\frac{1-y}{1+y}}\right), & y < 1 \\ \frac{1}{y} \left(\frac{94}{15}y(y^2-1)^2 - 3(1-2y^2)\right), & y = 1 \\ \left(\frac{8}{y^3} - \frac{20}{y} + 15y\right) \frac{2}{\sqrt{y^2-1}} \operatorname{arctan}\left(\sqrt{\frac{y-1}{y+1}}\right), & y > 1. \end{cases} \quad (\text{D9})$$

The radial flexion components are both zero.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.